b. Write the formula obtained by taking the derivative of both sides of the equation in part (a) with respect to x.

c. Use the result of part (b) to derive the formulas below.

(i) $2^{n-1} = \frac{1}{n}\left[\binom{n}{1} + 2\binom{n}{2} + 3\binom{n}{3} + \cdots + n\binom{n}{n}\right]$

(ii) $\sum_{k=1}^{n} k\binom{n}{k}(-1)^k = 0$

d. Express $\sum_{k=1}^{n} k\binom{n}{k}3^k$ in closed form (without using a summation sign or $\cdots$).

6.8 Probability Axioms and Expected Value

The theory of probability is at bottom nothing but common sense reduced to a calculus.
— Pierre-Simon Laplace (1749–1827)


Yevgeny Khaldei/CORBIS

Andrei Nikolaevich Kolmogorov (1903–1987)

Up to this point, you have calculated probabilities only for situations, such as tossing a fair coin or rolling a pair of balanced dice, where the outcomes in the sample space are all equally likely. But coins are not always fair and dice are not always balanced. How is it possible to calculate probabilities for these more general situations?

The following axioms were formulated by A. N. Kolmogorov in 1933 to provide a theoretical foundation for a far-ranging theory of probability. In this section we state the axioms, derive a few consequences, and introduce the notion of expected value.

Recall that a sample space is a set of all outcomes of a random process or experiment and that an event is a subset of a sample space.

Probability Axioms

Let S be a sample space, and let A and B be any events in S. Then

1. $0 \le P(A) \le 1$
2. $P(\emptyset) = 0$ and $P(S) = 1$
3. If A and B are disjoint (that is, if $A \cap B = \emptyset$), then the probability of the union of A and B is

$$P(A \cup B) = P(A) + P(B).$$

Example 6.8.1 Applying the Probability Axioms

Suppose that A and B are events in a sample space S. If A and B are mutually disjoint, could $P(A) = 0.6$ and $P(B) = 0.8$?

Solution No. Probability axiom 3 would imply that $P(A \cup B) = P(A) + P(B) = 0.6 + 0.8 = 1.4$, and since $1.4 > 1$, this result would violate probability axiom 1. ■

Example 6.8.2 The Probability of the Complement of an Event

Suppose that A is an event in a sample space S. Deduce that $P(A^c) = 1 - P(A)$.

Solution By Theorem 5.2.2(5), with S playing the role of the universal set U,

$$A \cap A^c = \emptyset \quad \text{and} \quad A \cup A^c = S.$$

Thus S is the disjoint union of A and A^c, and so

$$P(A \cup A^c) = P(A) + P(A^c) = P(S) = 1.$$

Subtracting $P(A)$ from both sides gives the result that $P(A^c) = 1 - P(A)$. ■

Probability of the Complement of an Event

If A is any event in a sample space S, then

$$P(A^c) = 1 - P(A). \tag{6.8.1}$$

It is important to check that Kolmogorov's probability axioms are consistent with the results obtained using the equally likely probability formula. To see that this is the case, let S be a finite sample space with outcomes $a_1, a_2, a_3, \ldots, a_n$. It is clear that all the singleton sets $\{a_1\}, \{a_2\}, \{a_3\}, \ldots, \{a_n\}$ are mutually disjoint and that their union is S. Since $P(S) = 1$, probability axiom 3 can be applied multiple times (see exercise 13 at the end of this section) to obtain

$$P(\{a_1\} \cup \{a_2\} \cup \{a_3\} \cup \cdots \cup \{a_n\}) = \sum_{k=1}^{n} P(\{a_k\}) = 1.$$

If, in addition, all the outcomes are equally likely, there is a positive real number c so that

$$P(\{a_1\}) = P(\{a_2\}) = P(\{a_3\}) = \cdots = P(\{a_n\}) = c.$$

Hence

$$1 = \sum_{k=1}^{n} c = \underbrace{c + c + \cdots + c}_{n \text{ terms}} = nc,$$

and thus

$$c = \frac{1}{n}.$$

It follows that if A is any event with outcomes $a_{i_1}, a_{i_2}, a_{i_3}, \ldots, a_{i_m}$, then

$$P(A) = \sum_{k=1}^{m} P(\{a_{i_k}\}) = \sum_{k=1}^{m} \frac{1}{n} = \frac{m}{n} = \frac{N(A)}{N(S)},$$

which is the result given by the equally likely probability formula.

Example 6.8.3 The Probability of a General Union of Two Events

Follow the steps outlined in parts (a) and (b) below to prove the following formula:

Probability of a General Union of Two Events

If S is any sample space and A and B are any events in S, then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \tag{6.8.2}$$

In both steps, suppose that A and B are any events in a sample space S.

a. Show that $A \cup B$ is a disjoint union of the following sets: $A - (A \cap B)$, $B - (A \cap B)$, and $A \cap B$.

b. In exercise 12 at the end of the section, you are asked to prove that for any events U and V in a sample space S, if $U \subseteq V$ then $P(V - U) = P(V) - P(U)$. Use this result and the result of part (a) to finish the proof of the formula.

Solution

a. Refer to Figure 6.8.1 as you read the following explanation. Elements in the set $A - (A \cap B)$ are in the region shaded blue, elements in $B - (A \cap B)$ are in the region shaded gray, and elements in $A \cap B$ are in the white region.

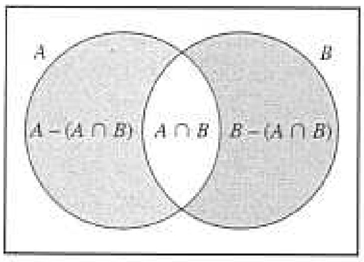


Figure 6.8.1

Part 1: Show that $A \cup B \subseteq (A - (A \cap B)) \cup (B - (A \cap B)) \cup (A \cap B)$: Given any element x in $A \cup B$, x satisfies exactly one of the following three conditions:

(1) $x \in A$ and $x \in B$

(2) $x \in A$ and $x \notin B$

(3) $x \in B$ and $x \notin A$

1. In the first case, $x \in A \cap B$, and so $x \in (A - (A \cap B)) \cup (B - (A \cap B)) \cup (A \cap B)$ by definition of union.
2. In the second case, $x \notin A \cap B$ (because $x \notin B$), and so $x \in A - (A \cap B)$. Therefore $x \in (A - (A \cap B)) \cup (B - (A \cap B)) \cup (A \cap B)$ by definition of union.
3. In the third case, $x \notin A \cap B$ (because $x \notin A$), and hence $x \in B - (A \cap B)$. So, again, $x \in (A - (A \cap B)) \cup (B - (A \cap B)) \cup (A \cap B)$ by definition of union.

Hence, in all three cases, $x \in (A - (A \cap B)) \cup (B - (A \cap B)) \cup (A \cap B)$, which completes the proof of part 1.

Moreover, since the three conditions are mutually exclusive, the three sets $A - (A \cap B)$, $B - (A \cap B)$, and $A \cap B$ are mutually disjoint.

Part 2: Show that $(A - (A \cap B)) \cup (B - (A \cap B)) \cup (A \cap B) \subseteq A \cup B$: Suppose x is any element in $(A - (A \cap B)) \cup (B - (A \cap B)) \cup (A \cap B)$. By definition of union, $x \in A - (A \cap B)$ or $x \in B - (A \cap B)$ or $x \in A \cap B$.

1. In case $x \in A - (A \cap B)$, then $x \in A$ and $x \notin A \cap B$ by definition of set difference. In particular, $x \in A$ and so $x \in A \cup B$.
2. In case $x \in B - (A \cap B)$, then $x \in B$ and $x \notin A \cap B$ by definition of set difference. In particular, $x \in B$ and so $x \in A \cup B$.
3. In case $x \in A \cap B$, then in particular, $x \in A$ and so $x \in A \cup B$.

Hence, in all three cases, $x \in A \cup B$, which completes the proof of part 2.

b. $P(A \cup B) = P((A - (A \cap B)) \cup (B - (A \cap B)) \cup (A \cap B))$ by part (a)

$= P(A - (A \cap B)) + P(B - (A \cap B)) + P(A \cap B)$
by exercise 13 at the end of the section and the fact that $A - (A \cap B)$, $B - (A \cap B)$, and $A \cap B$ are mutually disjoint

$= P(A) - P(A \cap B) + P(B) - P(A \cap B) + P(A \cap B)$
by exercise 12 at the end of the section because $A \cap B \subseteq A$ and $A \cap B \subseteq B$

$= P(A) + P(B) - P(A \cap B)$ by algebra. ■

Example 6.8.4 Computing the Probability of a General Union of Two Events

Suppose a card is chosen at random from an ordinary 52-card deck (see Section 6.1). What is the probability that the card is a face card (jack, queen, or king) or is from one of the red suits (hearts or diamonds)?

Solution Let A be the event that the chosen card is a face card, and let B be the event that the chosen card is from one of the red suits. The event that the card is a face card or is from one of the red suits is $A \cup B$. Now $N(A) = 4 \cdot 3 = 12$ (because each of the four suits has three face cards), and so $P(A) = 12/52$. Also $N(B) = 26$ (because half the cards are red), and so $P(B) = 26/52$. Finally, $N(A \cap B) = 6$ (because there are three face cards in hearts and another three in diamonds), and so $P(A \cap B) = 6/52$. It follows from the formula for the probability of a union of any two events that

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{12}{52} + \frac{26}{52} - \frac{6}{52} = \frac{32}{52} \cong 61.5\%.$$

Thus the probability that the chosen card is a face card or is from one of the red suits is approximately 61.5%. ■

Expected Value

People who buy lottery tickets regularly often justify the practice by saying that, even though they know that on average they will lose money, they are hoping for one significant gain, after which they believe they will quit playing. Unfortunately, when people who have lost money on a string of losing lottery tickets win some or all of it back, they generally decide to keep trying their luck instead of quitting.

The technical way to say that on average a person will lose money on the lottery is to say that the *expected value* of playing the lottery is negative.

• Definition

Suppose the possible outcomes of an experiment, or random process, are real numbers $a_1, a_2, a_3, \ldots, a_n$, which occur with probabilities $p_1, p_2, p_3, \ldots, p_n$. The **expected value** of the process is

$$\sum_{k=1}^{n} a_k p_k = a_1 p_1 + a_2 p_2 + a_3 p_3 + \cdots + a_n p_n.$$

Example 6.8.5 Expected Value of a Lottery

Suppose that 500,000 people pay \$5 each to play a lottery game with the following prizes: a grand prize of \$1,000,000, 10 second prizes of \$1,000 each, 1,000 third prizes of \$500 each, and 10,000 fourth prizes of \$10 each. What is the expected value of the game?

Solution Each of the 500,000 people has the same chance as any other of picking a winning lottery number, and so $p_k = \dfrac{1}{500000}$ for all $k = 1, 2, 3, \ldots, 500000$. Let $a_1, a_2, a_3, \ldots, a_{500000}$ be the net gains of the people playing the lottery. Let $a_1 = 999995$ (the net gain for the grand prize winner, which is one million dollars minus the \$5 cost of the winning ticket), $a_2 = a_3 = \cdots = a_{11} = 995$ (the net gain for each of the 10 second prize winners), $a_{12} = a_{13} = \cdots = a_{1011} = 495$ (the net gain for each of the 1,000 third prize winners), and $a_{1012} = a_{1013} = \cdots = a_{11011} = 5$ (the net gain for each of the 10,000 fourth prize winners). Since the remaining 488,989 people just lose their \$5, $a_{11012} = a_{11013} = \cdots = a_{500000} = -5$. The expected value of the game is therefore

$$\begin{aligned}
\sum_{k=1}^{500000} a_k p_k &= \sum_{k=1}^{500000}\left(a_k \cdot \frac{1}{500000}\right) && \text{because each } p_k = 1/500000\\
&= \frac{1}{500000}\sum_{k=1}^{500000} a_k && \text{by Theorem 4.1.1(2)}\\
&= \frac{1}{500000}(999995 + 10\cdot 995 + 1000\cdot 495 + 10000\cdot 5 + (-5)\cdot 488989)\\
&= \frac{1}{500000}(999995 + 9950 + 495000 + 50000 - 2444945)\\
&= -1.78.
\end{aligned}$$

In other words, a person who continues to play this lottery for a very long time will probably win some money occasionally but on average will lose \$1.78 per game. ■

Exercise Set 6.8

1. In any sample space S, what is $P(\emptyset)$?

2. Suppose A, B, and C are mutually exclusive events in a sample space S, $A \cup B \cup C = S$, and A and B have probabilities 0.3 and 0.5, respectively.
   a. What is $P(A \cup B)$? b. What is $P(C)$?

3. Suppose A and B are mutually exclusive events in a sample space S, C is another event in S, $A \cup B \cup C = S$, and A and B have probabilities 0.4 and 0.2 respectively.
   a. What is $P(A \cup B)$?
   b. Is it possible that $P(C) = 0.2$? Explain.

4. Suppose A and B are events in a sample space S with probabilities 0.8 and 0.7, respectively. Suppose also that $P(A \cap B) = 0.6$. What is $P(A \cup B)$?

5. Suppose A and B are events in a sample space S and suppose that $P(A) = 0.6$, $P(B^c) = 0.4$, and $P(A \cap B) = 0.2$. What is $P(A \cup B)$?

6. Suppose U and V are events in a sample space S and suppose that $P(U^c) = 0.3$, $P(V) = 0.6$, and $P(U^c \cup V^c) = 0.4$. What is $P(U \cup V)$?

7. Suppose a sample space S consists of three outcomes: 0, 1, and 2. Let $A = \{0\}$, $B = \{1\}$, and $C = \{2\}$, and suppose $P(A) = 0.4$, and $P(B) = 0.3$. Find each of the following:
   a. $P(A \cup B)$ b. $P(C)$ c. $P(A \cup C)$
   d. $P(A^c)$ e. $P(A^c \cap B^c)$ f. $P(A^c \cup B^c)$

8. Redo exercise 7 assuming that $P(A) = 0.5$ and $P(B) = 0.4$.

9. Let A and B be events in a sample space S, and let $C = S - (A \cup B)$. Suppose $P(A) = 0.4$, $P(B) = 0.5$, and $P(A \cap B) = 0.2$. Find each of the following:
   a. $P(A \cup B)$ b. $P(C)$ c. $P(A^c)$
   d. $P(A^c \cap B^c)$ e. $P(A^c \cup B^c)$ f. $P(B^c \cap C)$

10. Redo exercise 9 assuming that $P(A) = 0.7$, $P(B) = 0.3$, and $P(A \cap B) = 0.1$.

H 11. Prove that if S is any sample space and U and V are events in S with $U \subseteq V$, then $P(U) \le P(V)$.

H 12. Prove that if S is any sample space and U and V are any events in S, then $P(V - U) = P(V) - P(U \cap V)$.

H **13.** Use the axioms for probability and mathematical induction to prove that for all integers $n \geq 2$, if $A_1, A_2, A_3, \ldots, A_n$ are any mutually disjoint events in a sample space S, then

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_n) = \sum_{k=1}^{n} P(A_k).$$

14. A lottery game offers \$2 million to the grand prize winner, \$20 to each of 10,000 second prize winners, and \$4 to each of 50,000 third prize winners. The cost of the lottery is \$2 per ticket. Suppose that 1.5 million tickets are sold. What is the expected gain or loss of a ticket?

15. A company sends millions of people an entry form for a sweepstakes accompanied by an order form for magazine subscriptions. The first, second, and third prizes are \$10,000,000, \$1,000,000, and \$50,000, respectively. In order to qualify for a prize, a person is not required to order any magazines but has to spend 60 cents to mail back the entry form. If 30 million people qualify by sending back their entry forms, what is a person's expected gain or loss?

16. An urn contains four balls numbered 2, 2, 5, and 6. If a person selects a set of two balls at random, what is the expected value of the sum of the numbers on the balls?

17. An urn contains five balls numbered 1, 2, 2, 8, and 8. If a person selects a set of two balls at random, what is the expected value of the sum of the numbers on the balls?

18. An urn contains five balls numbered 1, 2, 2, 8, and 8. If a person selects a set of three balls at random, what is the expected value of the sum of the numbers on the balls?

19. When a pair of balanced dice are rolled and the sum of the numbers showing face up is computed, the result can be any number from 2 to 12, inclusive. What is the expected value of the sum?

H **20.** Suppose a person offers to play a game with you. In this game, when you draw a card from a standard 52-card deck, if the card is a face card you win \$3, and if the card is anything else you lose \$1. If you agree to play the game, what is your expected gain or loss?

21. A person pays \$1 to play the following game: The person tosses a fair coin four times. If no heads occur, the person pays an additional \$2, if one head occurs, the person pays an additional \$1, if two heads occur, the person just loses the initial dollar, if three heads occur, the person wins \$3, and if four heads occur, the person wins \$4. What is the person's expected gain or loss?

H **22.** A fair coin is tossed until either a head comes up or four tails are obtained. What is the expected number of tosses?

6.9 Conditional Probability, Bayes' Formula, and Independent Events

It is remarkable that a science which began with the consideration of games of chance should have become the most important object of human knowledge . . . . The most important questions of life are, for the most part, really only problems of probability.
— Pierre-Simon Laplace 1749–1827

In this section we introduce the notion of conditional probability and discuss Bayes' theorem and the kind of interesting results to which it leads. We then define the concept of independent events and give some applications.

Conditional Probability

Imagine a couple with two children, each of whom is equally likely to be a boy or a girl. Now suppose you are given the information that one is a boy. What is the probability that the other child is a boy?

Figure 6.9.1 shows the four equally likely combinations of gender for the children. You can imagine that the first letter refers to the older child and the second letter to the younger. Thus the combination BG indicates that the older child is a boy and the younger is a girl.

Figure 6.9.1

The combinations where one of the children is a boy are shaded gray, and the combination where the other child is also a boy is shaded blue. Given that you know one child is a boy, only the three combinations in the gray region could be the case, so you can think of the set of those outcomes as a new sample space with three elements, all of which are equally likely. Within the new sample space, there is one combination where the other child is a boy (in the region shaded blue-gray). Thus it would be reasonable to say that the likelihood that the other child is a boy, given that at least one is a boy, is $1/3 = 33\frac{1}{3}\%$. Note that because the original sample space contained four outcomes,

$$\frac{P(\text{at least one child is a boy and the other child is also a boy})}{P(\text{at least one child is a boy})} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$$

also. A generalization of this observation forms the basis for the following definition.

• Definition

Let A and B be events in a sample space S. If $P(A) \neq 0$, then the **conditional probability of B given A, denoted $P(B \mid A)$,** is

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}. \qquad 6.9.1$$

Example 6.9.1 Computing a Conditional Probability

A pair of fair dice, one blue and the other gray, are rolled. What is the probability that the sum of the numbers showing face up is 8, given that both of the numbers are even?

Solution The sample space is the set of all 36 outcomes obtained from rolling the two dice and noting the numbers showing face up on each. As in Section 6.1, denote by ab the outcome that the number showing face up on the blue die is a and the one on the gray die is b. Let A be the event that both numbers are even and B the event that the sum of the numbers is 8. Then $A = \{22, 24, 26, 42, 44, 46, 62, 64, 66\}$, $B = \{26, 35, 44, 53, 62\}$, and $A \cap B = \{26, 44, 62\}$. Because the dice are fair (so all outcomes are equally likely), $P(A) = 9/36$, $P(B) = 5/36$ and $P(A \cap B) = 3/36$. By definition of conditional probability,

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{3}{36}}{\frac{9}{36}} = \frac{3}{9} = \frac{1}{3}.$$ ■

Note that when both sides of the formula for conditional probability (formula 6.9.1) are multiplied by $P(A)$, a formula for $P(A \cap B)$ is obtained:

$$P(A \cap B) = P(B \mid A) \cdot P(A). \qquad 6.9.2$$

Dividing both sides of formula (6.9.2) by $P(B \mid A)$ gives a formula for $P(A)$:

$$P(A) = \frac{P(A \cap B)}{P(B \mid A)}. \qquad 6.9.3$$

Example 6.9.2 Further Applications of the Conditional Probability Formula

An urn contains 5 blue and 7 gray balls. Let us say that 2 are chosen at random, one after the other, without replacement.

a. What is the probability that both balls are blue?

b. What is the probability that the second ball is blue but the first ball is not?

c. What is the probability that the second ball is blue?

d. What is the probability that at least one of the balls is blue?

e. If the experiment of choosing 2 balls from the urn were repeated many times over, what would be the expected value of the number of blue balls?

Solution Let S denote the sample space of all possible choices of two balls from the urn, let E be the event that the first ball is blue, and let F be the event that the second ball is blue.

a. The probability that both balls are blue is $P(E \cap F)$. Because there are 12 balls of which 5 are blue, the probability that the first ball is blue is

$$P(E) = \frac{5}{12}.$$

If the first ball is blue, then when the second ball is chosen the urn will contain 4 blue and 7 gray balls. Thus $P(F \mid E) = 4/11$, so by formula (6.9.2),

$$P(E \cap F) = P(F \mid E) \cdot P(E) = \frac{4}{11} \cdot \frac{5}{12} = \frac{20}{132}.$$

b. The probability that the second ball is blue but the first ball is not is $P(F \cap E^c)$. To compute this number, note that because there are 12 balls of which 7 are not blue, $P(E^c) = 7/12$. Also if the first ball is not blue, then when the second ball is chosen, the urn will contain 5 blue and 6 gray balls, and thus $P(F \mid E^c) = 5/11$. Hence, by formula 6.9.2,

$$P(F \cap E^c) = P(F \mid E^c) \cdot P(E^c) = \frac{5}{11} \cdot \frac{7}{12} = \frac{35}{132}.$$

c. The event that the second ball is blue can occur in one of two mutually exclusive ways: Either the first ball is blue and the second is also blue, or the first ball is gray and the second is blue. In other words, F is the disjoint union of $F \cap E$ and $F \cap E^c$. Hence

$$\begin{aligned}
P(F) &= P((F \cap E) \cup (F \cap E^c)) \\
&= P(F \cap E) + P(F \cap E^c) && \text{by probability axiom 3} \\
&= \frac{20}{132} + \frac{35}{132} && \text{by parts (a) and (b)} \\
&= \frac{55}{132} = \frac{5}{12}.
\end{aligned}$$

Thus the probability that the second ball is blue is $5/12$, the same as the probability that the first ball is blue.

d. By formula 6.8.2, for the union of any two events,

$$\begin{aligned} P(E \cup F) &= P(E) + P(F) - P(E \cap F) \\ &= \frac{5}{12} + \frac{5}{12} - \frac{20}{132} \qquad \text{by parts (a) and (c)} \\ &= \frac{90}{132} = \frac{15}{22}. \end{aligned}$$

Thus the probability is 15/22, or approximately 68.2%, that at least one of the balls is blue.

e. The event that neither ball is blue is the complement of the event that at least one of the balls is blue, so

$$\begin{aligned} P(0 \text{ blue balls}) &= 1 - P(\text{at least one ball is blue}) \qquad \text{by formula 6.8.1} \\ &= 1 - \frac{15}{22} \qquad \text{by part (d)} \\ &= \frac{7}{22}. \end{aligned}$$

The event that one ball is blue can occur in one of two mutually exclusive ways: Either the second ball is blue and the first is not, or the first ball is blue and the second is not. Part (b) showed that the probability of the first way is $\frac{35}{132}$, and the same technique shows that the probability of the second way is also $\frac{35}{132}$. Thus, by probability axiom 3,

$$P(1 \text{ blue ball}) = \frac{35}{132} + \frac{35}{132} = \frac{70}{132}.$$

Finally, by part (a),

$$P(2 \text{ blue balls}) = \frac{20}{132}.$$

Therefore,

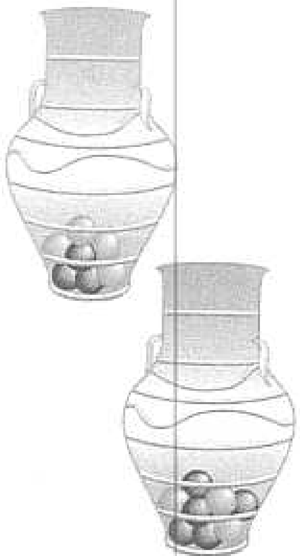
$$\begin{aligned} \begin{bmatrix} \text{the expected value of} \\ \text{the number of blue balls} \end{bmatrix} &= 0 \cdot P(0 \text{ blue balls}) + 1 \cdot P(1 \text{ blue ball}) \\ &\qquad + 2 \cdot P(2 \text{ blue balls}) \\ &= 0 \cdot \frac{7}{22} + 1 \cdot \frac{70}{132} + 2 \cdot \frac{20}{132} \\ &= \frac{110}{132} \cong 0.8. \end{aligned}$$

■

Bayes' Theorem

Suppose that one urn contains 3 blue and 4 gray balls and a second urn contains 5 blue and 3 gray balls. A ball is selected by choosing one of the urns at random and then picking a ball at random from that urn. If the chosen ball is blue, what is the probability that it came from the first urn?

This problem can be solved by carefully interpreting all the information that is known and putting it together in just the right way. Let A be the event that the chosen ball is blue, B_1 the event that the ball came from the first urn, and B_2 the event that the ball came from



the second urn. Because 3 of the 7 balls in urn one are blue, and 5 of the 8 balls in urn two are blue,

$$P(A \mid B_1) = \frac{3}{7} \quad \text{and} \quad P(A \mid B_2) = \frac{5}{8}.$$

And because the urns are equally likely to be chosen,

$$P(B_1) = P(B_2) = \frac{1}{2}.$$

Moreover, by formula (6.9.2),

$$P(A \cap B_1) = P(A \mid B_1) \cdot P(B_1) = \frac{3}{7} \cdot \frac{1}{2} = \frac{3}{14}, \quad \text{and}$$

$$P(A \cap B_2) = P(A \mid B_2) \cdot P(B_2) = \frac{5}{8} \cdot \frac{1}{2} = \frac{5}{16}.$$

But A is the disjoint union of $(A \cap B_1)$ and $(A \cap B_2)$, so by probability axiom 3,

$$P(A) = P((A \cap B_1) \cup (A \cap B_2)) = P(A \cap B_1) + P(A \cap B_2) = \frac{3}{14} + \frac{5}{16} = \frac{59}{112}.$$

Finally, by definition of conditional probability,

$$P(B_1 \mid A) = \frac{P(B_1 \cap A)}{P(A)} = \frac{\frac{3}{14}}{\frac{59}{112}} = \frac{336}{826} \cong 40.7\%.$$

Thus, if the chosen ball is blue, the probability is approximately 40.7% that it came from the first urn.

Thomas Bayes (1702–1761)

The steps used to derive the answer in the example above can be generalized to prove Bayes' theorem. (See exercises 8 and 9 at the end of this section.) Thomas Bayes was an English Presbyterian minister who devoted much of his energies to mathematics. The theorem that bears his name was published posthumously in 1763. The portrait at the left is the only one attributed to him, but its authenticity has recently come into question.

Theorem 6.9.1 Bayes' Theorem

Suppose that a sample space S is a union of mutually disjoint events $B_1, B_2, B_3, \ldots, B_n$, suppose A is an event in S, and suppose A and all the B_i have nonzero probabilities. If k is an integer with $1 \le k \le n$, then

$$P(B_k \mid A) = \frac{P(A \mid B_k) \cdot P(B_k)}{P(A \mid B_1) \cdot P(B_1) + P(A \mid B_2) \cdot P(B_2) + \cdots + P(A \mid B_n) \cdot P(B_n)}$$

Example 6.9.3 Applying Bayes' Theorem

Most medical tests occasionally produce incorrect results, called false positives and false negatives. When a test is designed to determine whether a patient has a certain disease, a **false positive** result indicates that a patient has the disease when the patient does not have it. A **false negative** result indicates that a patient does not have the disease when the patient does have it.

When large-scale health screenings are performed for diseases with relatively low incidence, those who develop the screening procedures have to balance several considerations: the per-person cost of the screening, follow-up costs for further testing of false positives, and the possibility that people who have the disease will develop unwarranted confidence in the state of their health.

Consider a medical test that screens for a disease found in 5 people in 1,000. Suppose that the false positive rate is 3% and the false negative rate is 1%. Then 99% of the time a person who has the condition tests positive for it, and 97% of the time a person who does not have the condition tests negative for it. (See exercise 3 at the end of this section.)

a. What is the probability that a randomly chosen person who tests positive for the disease actually has the disease?

b. What is the probability that a randomly chosen person who tests negative for the disease does not indeed have the disease?

Solution Consider a person chosen at random from among those screened. Let A be the event that the person tests positive for the disease, B_1 the event that the person actually has the disease, and B_2 the event that the person does not have the disease. Then

$$P(A \mid B_1) = 0.99, \quad P(A^c \mid B_1) = 0.01, \quad P(A^c \mid B_2) = 0.97, \quad \text{and} \quad P(A \mid B_2) = 0.03.$$

Also, because 5 people in 1,000 have the disease,

$$P(B_1) = 0.005 \quad \text{and} \quad P(B_2) = 0.995.$$

a. By Bayes' theorem,

$$\begin{aligned} P(B_1 \mid A) &= \frac{P(A \mid B_1) \cdot P(B_1)}{P(A \mid B_1) \cdot P(B_1) + P(A \mid B_2) \cdot P(B_2)} \\ &= \frac{(0.99) \cdot (0.005)}{(0.99) \cdot (0.005) + (0.03) \cdot (0.995)} \\ &\cong 0.1422 \cong 14.2\%. \end{aligned}$$

Thus the probability that a person with a positive test result actually has the disease is approximately 14.2%.

b. By Bayes' theorem,

$$\begin{aligned} P(B_2 \mid A^c) &= \frac{P(A^c \mid B_2) \cdot P(B_2)}{P(A^c \mid B_1) \cdot P(B_1) + P(A^c \mid B_2) \cdot P(B_2)} \\ &= \frac{(0.97) \cdot (0.995)}{(0.01) \cdot (0.005) + (0.97) \cdot (0.995)} \\ &\cong 0.999948 \cong 99.995\%. \end{aligned}$$

Thus the probability that a person with a negative test result does not have the disease is approximately 99.995%.

You might be surprised by these numbers, but they are fairly typical of the situation where the screening test is significantly less expensive than a more accurate test for the same disease yet produces positive results for nearly all people with the disease. Using the screening test limits the expense of unnecessarily using the more costly test to a relatively small percentage of the population being screened, while only rarely indicating that a person who has the disease is free of it. ■

Independent Events

Suppose a coin is tossed twice. It seems intuitively clear that the outcome of the first toss does not depend in any way on the outcome of the second toss, and conversely. In other words, if, for instance, A is the event that a head is obtained on the first toss and B is the event that a head is obtained on the second toss, then if the coin is tossed randomly both times, events A and B should be *independent* in the sense that $P(A \mid B) = P(A)$ and $P(B \mid A) = P(B)$. This intuitive idea of independence is supported by the following analysis. If the coin is fair, then the four outcomes HH, HT, TH, and TT are equally likely, and

$$A = \{HH, HT\}, \quad B = \{TH, HH\}, \quad A \cap B = \{HH\}.$$

Hence

$$P(A) = P(B) = \frac{2}{4} = \frac{1}{2}.$$

But also

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2} \quad \text{and} \quad P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2},$$

and thus $P(A \mid B) = P(A)$ and $P(B \mid A) = P(B)$.

To obtain the final form for definition of independence, observe that

if $P(B) \neq 0$ and $P(A \mid B) = P(A)$, then $P(A \cap B) = P(A \mid B) \cdot P(B) = P(A) \cdot P(B)$.

By the same argument,

if $P(A) \neq 0$ and $P(B \mid A) = P(B)$, then $P(A \cap B) = P(A) \cdot P(B)$.

Conversely (see exercise 17 at the end of this section),

if $P(A \cap B) = P(A) \cdot P(B)$ and $P(A) \neq 0$, then $P(B \mid A) = P(B)$,

and

if $P(A \cap B) = P(A) \cdot P(B)$ and $P(B) \neq 0$, then $P(A \mid B) = P(A)$.

Thus, for convenience and to eliminate the requirement that the probabilities be nonzero, we use the following product formula to define independent events.

• Definition

If A and B are events in a sample space S, then A and B are **independent** if, and only if,

$$P(A \cap B) = P(A) \cdot P(B).$$

It would be natural to think that mutually disjoint events would be independent, but in fact almost the opposite is true: Mutually disjoint events with nonzero probabilities are dependent.

Example 6.9.4 Disjoint Events and Independence

Let A and B be events in a sample space S, and suppose $A \cap B = \emptyset$, $P(A) \neq 0$, and $P(B) \neq 0$. Show that $P(A \cap B) \neq P(A) \cdot P(B)$.

Solution Because $A \cap B = \emptyset$, $P(A \cap B) = 0$ by probability axiom 2. But $P(A) \cdot P(B) \neq 0$ because neither $P(A)$ nor $P(B)$ equals zero. Thus $P(A \cap B) \neq P(A) \cdot P(B)$. ■

The following example, and its immediate consequence, show how the independence of two events extends to their complements.

Example 6.9.5 The Probability of $A \cap B^c$ When A and B Are Independent Events

Suppose A and B are independent events in a sample space S. Show that A and B^c are also independent.

Solution Observe that

$$\begin{aligned}(A \cap B) \cup (A \cap B^c) &= A \cap (B \cup B^c) && \text{by the distributive law for sets}\\ &= A \cap S && \text{by the complement law for union}\\ &= A && \text{by the identity law for intersection}\end{aligned}$$

Also $(A \cap B) \cap (A \cap B^c) = \emptyset$ because $B \cap B^c = \emptyset$. Apply probability axiom 3 to the above equality to obtain

$$P((A \cap B) \cup (A \cap B^c)) = P(A \cap B) + P(A \cap B^c) = P(A).$$

Solving for $P(A \cap B^c)$ gives that

$$\begin{aligned}P(A \cap B^c) &= P(A) - P(A \cap B)\\ &= P(A) - P(A) \cdot P(B) && \text{because } A \text{ and } B \text{ are independent}\\ &= P(A)(1 - P(B)) && \text{by factoring out } P(A)\\ &= P(A) \cdot P(B^c) && \text{by formula 6.8.1.}\end{aligned}$$

Thus A and B^c are independent events. ■

It follows immediately from Example 6.9.5 that if A and B are independent, then A^c and B are also independent and so are A^c and B^c. (See exercises 21 and 22 at the end of this section.) These results are applied in Example 6.9.6.

Example 6.9.6 Computing Probabilities of Intersections of Independent Events

A coin is loaded so that the probability of heads is 0.6. Suppose the coin is tossed twice. Although the probability of heads is greater than the probability of tails, there is no reason to believe that whether the coin lands heads or tails on one toss will affect whether it lands heads or tails on the other toss. Thus it is reasonable to assume that the results of the tosses are independent.

a. What is the probability of obtaining two heads?

b. What is the probability of obtaining one head?

c. What is the probability of obtaining no heads?

d. What is the probability of obtaining at least one head?

Solution The sample space S consists of the four outcomes $\{HH, HT, TH, TT\}$, which are not equally likely. Let E be the event that a head is obtained on the first toss, and let F be the event that a head is obtained on the second toss. Then $P(E) = P(F) = 0.6$, and it is to be assumed that E and F are independent.

a. The probability of obtaining two heads is $P(E \cap F)$. Because E and F are independent,

$$P(\text{two heads}) = P(E \cap F) = P(E) \cdot P(F) = (0.6)(0.6) = 0.36 = 36\%.$$

b. One head can be obtained in two mutually exclusive ways: head on the first toss and tail on the second, or tail on the first toss and head on the second. Thus, the event of obtaining exactly one head is $(E \cap F^c) \cup (E^c \cap F)$. Also $(E \cap F^c) \cap (E^c \cap F) = \emptyset$, and, moreover, by the formula for the probability of the complement of an event, $P(E^c) = P(F^c) = 1 - 0.6 = 0.4$. Hence

$$\begin{aligned} P(\text{one head}) &= P((E \cap F^c) \cup (E^c \cap F)) \\ &= P(E) \cdot P(F^c) + P(E^c) \cdot P(F) \quad \text{by Example 6.9.5 and exercise 21} \\ &= (0.6)(0.4) + (0.4)(0.6) \\ &= 0.48 = 48\%. \end{aligned}$$

c. The probability of obtaining no heads is $P(E^c \cap F^c)$. By exercise 22,

$$P(\text{no heads}) = P(E^c \cap F^c) = P(E^c) \cdot P(F^c) = (0.4)(0.4) = 0.16 = 16\%.$$

d. There are two ways to solve this problem. One is to observe that because the event of obtaining one head and the event of obtaining two heads are mutually disjoint,

$$\begin{aligned} P(\text{at least one head}) &= P(\text{one head}) + P(\text{two heads}) \\ &= 0.48 + 0.36 \quad \text{by parts (a) and (b)} \\ &= 0.84 = 84\%. \end{aligned}$$

The second way is to use the fact that the event of obtaining at least one head is the complement of the event of obtaining no heads. So

$$\begin{aligned} P(\text{at least one head}) &= 1 - P(\text{no heads}) \\ &= 1 - 0.16 \quad \text{by part (c)} \\ &= 0.84 = 84\%. \end{aligned}$$

■

Example 6.9.7 Expected Value of Tossing a Loaded Coin Twice

Suppose that a coin is loaded so that the probability of heads is 0.6, and suppose the coin is tossed twice. If this experiment is repeated many times, what is the expected value of the number of heads?

Solution Think of the outcomes of the coin tossings as just 0, 1, or 2 heads. Example 6.9.6 showed that the probabilities of these outcomes are 0.16, 0.48, and 0.36, respectively. Thus, by definition of expected value, the

$$\text{expected number of heads} = 0 \cdot (0.16) + 1 \cdot (0.48) + 2 \cdot (0.36) = 1.2.$$

■

What if a loaded coin is tossed more than twice? Suppose it is tossed ten times, or a hundred times. What are the probabilities of various numbers of heads? To answer this question, it is necessary to expand the notion of independence to more than two events. For instance, we say three events A, B, and C are *pairwise independent* if, and only if,

$$P(A \cap B) = P(A) \cdot P(B),\ P(A \cap C) = P(A) \cdot P(C), \text{ and } P(B \cap C) = P(B) \cdot P(C).$$

The next example shows that events can be pairwise independent without satisfying the condition $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$. Conversely, they can satisfy the condition $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$ without being pairwise independent (see exercise 26 at the end of this section).

Example 6.9.8 Exploring Independence for Three Events

Suppose that a fair coin is tossed twice. Let A be the event that a head is obtained on the first toss, B the event that a head is obtained on the second toss, and C the event that either two heads or two tails are obtained. Show that A, B, and C are pairwise independent but do not satisfy the condition $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$.

Solution Because there are four equally likely outcomes—HH, HT, TH, and TT—it is clear that $P(A) = P(B) = P(C) = \frac{1}{2}$. You can also see that $A \cap B = \{HH\}$, $A \cap C = \{HH\}$, $B \cap C = \{HH\}$, and $A \cap B \cap C = \{HH\}$. Hence $P(A \cap B) = P(A \cap C) = P(B \cap C) = \frac{1}{4}$, and so $P(A \cap B) = P(A) \cdot P(B)$, $P(A \cap C) = P(A) \cdot P(C)$, and $P(B \cap C) = P(B) \cdot P(C)$. Thus A, B, and C are pairwise independent. But

$$P(A \cap B \cap C) = P(\{HH\}) = \frac{1}{4} \neq \left(\frac{1}{2}\right)^3 = P(A) \cdot P(B) \cdot P(C). \quad \blacksquare$$

Because of situations like that in Example 6.9.8, four conditions must be included in the definition of independence for three events.

• Definition

Let A, B, and C be events in a sample space S. A, B, and C are **pairwise independent** if, and only if, they satisfy conditions 1–3 below. They are **mutually independent** if, and only if, they satisfy all four conditions below.

1. $P(A \cap B) = P(A) \cdot P(B)$
2. $P(A \cap C) = P(A) \cdot P(C)$
3. $P(B \cap C) = P(B) \cdot P(C)$
4. $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$

The definition of mutual independence for any collection of n events with $n \geq 2$ generalizes the two definitions given previously.

• Definition

Events $A_1, A_2, A_3, \ldots, A_n$ in a sample space S are **mutually independent** if, and only if, the probability of the intersection of any subset of the events is the product of the probabilities of the events in the subset.

Example 6.9.9 Tossing a Loaded Coin Ten Times

A coin is loaded so that the probability of heads is 0.6 (and thus the probability of tails is 0.4). Suppose the coin is tossed ten times. As in Example 6.9.6, it is reasonable to assume that the results of the tosses are mutually independent.

a. What is the probability of obtaining eight heads?

b. What is the probability of obtaining at least eight heads?

Solution

a. For each $i = 1, 2, \ldots, 10$, let H_i be the event that a head is obtained on the ith toss, and let T_i be the event that a tail is obtained on the ith toss. Suppose that the eight heads occur on the first eight tosses and that the remaining two tosses are tails. This is the event $H_1 \cap H_2 \cap H_3 \cap H_4 \cap H_5 \cap H_6 \cap H_7 \cap H_8 \cap T_9 \cap T_{10}$. For simplicity, we denote it as $HHHHHHHHTT$. By definition of mutually independent events,

$$P(HHHHHHHHTT) = (0.6)^8(0.4)^2.$$

Because of the commutative law for multiplication, if the eight heads occur on any other of the ten tosses, the same number is obtained. For instance, if we denote the event $H_1 \cap H_2 \cap T_3 \cap H_4 \cap H_5 \cap H_6 \cap H_7 \cap H_8 \cap T_9 \cap H_{10}$ by $HHTHHHHHTH$, then

$$P(HHTHHHHHTH) = (0.6)^2(0.4)(0.6)^5(0.4)(0.6) = (0.6)^8(0.4)^2.$$

Now there are as many different ways to obtain eight heads in ten tosses as there are subsets of eight elements (the toss numbers on which heads are obtained) that can be chosen from a set of ten elements. This number is $\binom{10}{8}$. It follows that, because the different ways of obtaining eight heads are all mutually exclusive,

$$P(\text{eight heads}) = \binom{10}{8}(0.6)^8(0.4)^2.$$

b. By reasoning similar to that in part (a),

$$P(\text{nine heads}) = \begin{bmatrix}\text{the number of different}\\ \text{ways nine heads can be}\\ \text{obtained in ten tosses}\end{bmatrix} \cdot (0.6)^9(0.4)^1 = \binom{10}{9}(0.6)^9(0.4),$$

and

$$P(\text{ten heads}) = \begin{bmatrix}\text{the number of different}\\ \text{ways ten heads can be}\\ \text{obtained in ten tosses}\end{bmatrix} \cdot (0.6)^{10}(0.4)^0 = \binom{10}{10}(0.6)^{10}.$$

Because obtaining eight, obtaining nine, and obtaining ten heads are mutually disjoint events,

$$\begin{aligned} P(\text{at least eight heads}) &= P(\text{eight heads}) + P(\text{nine heads}) + P(\text{ten heads}) \\ &= \binom{10}{8}(0.6)^8(0.4)^2 + \binom{10}{9}(0.6)^9(0.4) + \binom{10}{10}(0.6)^{10} \\ &\cong 0.167 = 16.7\%. \end{aligned}$$

■

Note the occurrence of the binomial coefficients $\binom{n}{k}$ in solutions to problems like the one in Example 6.9.9. For that reason, probabilities of the form

$$\binom{n}{k}p^{n-k}(1-p)^k,$$

where $0 \le p \le 1$, are called **binomial probabilities.** Binomial probabilities occur in situations with multiple, mutually independent repetitions of a random process, with the same two possible outcomes that have the same probabilities on each repetition.

Exercise Set 6.9

1. Suppose $P(A \mid B) = 1/2$ and $P(A \cap B) = 1/6$. What is $P(B)$?

2. Suppose $P(X \mid Y) = 1/3$ and $P(Y) = 1/4$. What is $P(X \cap Y)$?

3. a. Prove that if A and B are any events in a sample space S, with $P(B) \neq 0$, then $P(A^c \mid B) = 1 - P(A \mid B)$.
   b. Explain how this result justifies the following statements: (1) If the probability of a false positive on a test for a condition is 4%, then there is a 96% probability that a person who does not have the condition will have a negative test result. (2) If the probability of a false negative on a test for a condition is 1%, then there is a 99% probability that a person who does have the condition will test positive for it.

H 4. Suppose that A and B are events in a sample space S and that $P(A)$, $P(B)$, and $P(A \mid B)$ are known. Derive a formula for $P(A \mid B^c)$.

5. An urn contains 25 red balls and 15 blue balls. Two are chosen at random, one after the other, without replacement.
   a. What is the probability that both balls are red?
   b. What is the probability that the second ball is red but the first ball is not?
   c. What is the probability that the second ball is red?
   d. What is the probability that at least one of the balls is red?

6. Redo exercise 5 assuming that the urn contains 30 red balls and 40 blue balls.

7. A pool of 10 semifinalists for a job consists of 7 men and 3 women. Because all are considered equally qualified, the names of two of the semifinalists are drawn, one after the other, at random, to become finalists for the job.
   a. What is the probability that both finalists are women?
   b. What is the probability that both finalists are men?
   H c. What is the probability that one finalist is a woman and the other is a man?

H 8. Prove Bayes' theorem for $n = 2$. That is, prove that if a sample space S is a union of mutually disjoint events B_1 and B_2, if A is an event in S with $P(A) \neq 0$, and if $k = 1$ or $k = 2$, then

$$P(B_k \mid A) = \frac{P(A \mid B_k) \cdot P(B_k)}{P(A \mid B_1) \cdot P(B_1) + P(A \mid B_2) \cdot P(B_2)}.$$

9. Prove the full version of Bayes' theorem.

10. One urn contains 12 blue balls and 7 white balls, and a second urn contains 8 blue balls and 19 white balls. An urn is selected at random, and a ball is chosen from the urn.
    a. What is the probability that the chosen ball is blue?
    b. If the chosen ball is blue, what is the probability that it came from the first urn?

11. Redo exercise 10 assuming that the first urn contains 4 blue balls and 16 white balls and the second urn contains 10 blue balls and 9 white balls.

H 12. One urn contains 10 red balls and 25 green balls, and a second urn contains 22 red balls and 15 green balls. A ball is chosen as follows: First an urn is selected by tossing a loaded coin with probability 0.4 of landing heads up and probability 0.6 of landing tails up. If the coin lands heads up, the first urn is chosen; otherwise, the second urn is chosen. Then a ball is picked at random from the chosen urn.
    a. What is the probability that the chosen ball is green?
    b. If the chosen ball is green, what is the probability that it was picked from the first urn?

13. A drug-screening test is used in a large population of people of whom 4% actually use drugs. Suppose that the false positive rate is 3% and the false negative rate is 2%. Thus a person who uses drugs tests positive for them 98% of the time, and a person who does not use drugs tests negative for them 97% of the time.
    a. What is the probability that a randomly chosen person who tests positive for drugs actually uses drugs?
    b. What is the probability that a randomly chosen person who tests negative for drugs does not use drugs?

14. Two different factories both produce a certain automobile part. The probability that a component from the first factory is defective is 2%, and the probability that a component from the second factory is defective is 5%. In a supply of 180 of the parts, 100 were obtained from the first factory and 80 from the second factory.
    a. What is the probability that a part chosen at random from the 180 is from the first factory?
    b. What is the probability that a part chosen at random from the 180 is from the second factory?
    c. What is the probability that a part chosen at random from the 180 is defective?
    d. If the chosen part is defective, what is the probability that it came from the first factory?

H 15. Three different suppliers—X, Y, and Z—provide produce for a grocery store. Twelve percent of produce from X is superior grade, 8% of produce from Y is superior grade and 15% of produce from Z is superior grade. The store obtains 20% of its produce from X, 45% from Y, and 35% from Z.
    a. If a piece of produce is purchased, what is the probability that it is superior grade?
    b. If a piece of produce in the store is superior grade, what is the probability that it is from X?

16. Prove that if A and B are events in a sample space S with the property that $P(A \mid B) = P(A)$ and $P(A) \neq 0$, then $P(B \mid A) = P(B)$.

17. Prove that if $P(A \cap B) = P(A) \cdot P(B)$, $P(A) \neq 0$, and $P(B) \neq 0$, then $P(A \mid B) = P(A)$ and $P(B \mid A) = P(B)$.

18. A pair of fair dice, one blue and the other gray, are rolled. Let A be the event that the number face up on the blue die is 2, and let B be the event that the number face up on the gray die is 4 or 5. Show that $P(A \mid B) = P(A)$ and $P(B \mid A) = P(B)$.

19. Suppose a fair coin is tossed three times. Let A be the event that a head appears on the first toss, and let B be the event that an even number of heads is obtained. Show that $P(A \mid B) = P(A)$ and $P(B \mid A) = P(B)$.

20. If A and B are events in a sample space S and $A \cap B = \emptyset$, what must be true in order for A and B to be independent? Explain.

21. Prove that if A and B are independent events in a sample space S, then A^c and B are also independent. Do not use the result of Example 6.9.5.

22. Prove that if A and B are independent events in a sample space S, then A^c and B^c are also independent.

23. A student taking a multiple-choice exam does not know the answers to two questions. All have five choices for the answer. For one of the two questions, the student can eliminate two answer choices as incorrect but has no idea about the other answer choices. For the other question, the student has no clue about the correct answer at all. Assume that whether the student chooses the correct answer on one of the questions does not affect whether the student chooses the correct answer on the other question.

a. What is the probability that the student will answer both questions correctly?

b. What is the probability that the student will answer exactly one of the questions correctly?

c. What is the probability that the student will answer neither question correctly?

24. A company uses two proofreaders X and Y to check a certain manuscript. X misses 12% of typographical errors and Y misses 15%. Assume that the proofreaders work independently.

a. What is the probability that a randomly chosen typographical error will be missed by both proofreaders?

b. If the manuscript contains 1,000 typographical errors, what number can be expected to be missed?

25. A coin is loaded so that the probability of heads is 0.7 and the probability of tails is 0.3. Suppose that the coin is tossed twice and that the results of the tosses are independent.

a. What is the probability of obtaining exactly two heads?

b. What is the probability of obtaining exactly one head?

c. What is the probability of obtaining no heads?

d. What is the probability of obtaining at least one head?

* 26. Describe a sample space and events A, B, and C, where $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$ but A, B, and C are not pairwise independent.

H **27.** The example used to introduce conditional probability described a family with two children each of whom was equally likely to be a boy or a girl. The example showed that if it is known that one child is a boy, the probability that the other child is a boy is 1/3. Now imagine the same kind of family—two children each of whom is equally likely to be a boy or a girl. Suppose you meet one of the children and see that it is a boy. What is the probability that the other child is a boy? Explain. (Be careful. The answer may surprise you.)

28. A coin is loaded so that the probability of heads is 0.7 and the probability of tails is 0.3. Suppose that the coin is tossed ten times and that the results of the tosses are mutually independent.

a. What is the probability of obtaining exactly seven heads?

b. What is the probability of obtaining exactly ten heads?

c. What is the probability of obtaining no heads?

d. What is the probability of obtaining at least one head?

29. Suppose that ten items are chosen at random from a large batch delivered to a company. The manufacturer claims that just 3% of the items in the batch are defective. Assume that the batch is large enough so that even though the selection is made without replacement, the number 0.03 can be used to approximate the probability that any one of the ten items is defective. In addition, assume that because the items are chosen at random, the outcomes of the choices are mutually independent. Finally, assume that the manufacturer's claim is correct.

a. What is the probability that none of the ten is defective?

b. What is the probability that at least one of the ten is defective?

c. What is the probability that exactly four of the ten are defective?

d. What is the probability that at most two of the ten are defective?

30. Suppose the probability of a false positive result on a mammogram is 4% and that radiologists' interpretations of mammograms are mutually independent in the sense that whether or not a radiologist finds a positive result on one mammogram does not influence whether or not the radiologist finds a positive result on another mammogram. Assume that a woman has a mammogram every year for ten years.

a. What is the probability that she will have no false positive results during that time?

b. What is the probability that she will have at least one false positive result during that time?

c. What is the probability that she will have exactly two false positive results during that time?

d. Suppose that the probability of a false negative result on a mammogram is 2%, and assume that the probability that a randomly chosen woman has breast cancer is 0.0002.

(i) If a woman has a positive test result one year, what is the probability that she actually has breast cancer?

(ii) If a woman has a negative test result one year, what is the probability that she actually has breast cancer?

31. Empirical data indicate that approximately 103 out of every 200 children born are male. Hence the probability of a newborn being male is about 51.5%. Suppose that a family has six children, and suppose that the genders of all the children are mutually independent.

H a. What is the probability that none of the children is male?
b. What is the probability that at least one of the children is male?
c. What is the probability that exactly five of the children are male?

32. A person takes a multiple-choice exam in which each question has four possible answers. Suppose that the person has no idea about the answers to three of the questions and simply chooses randomly for each one.
a. What is the probability that the person will answer all three questions correctly?
b. What is the probability that the person will answer exactly two questions correctly?
c. What is the probability that the person will answer exactly one question correctly?
d. What is the probability that the person will answer no questions correctly?
e. Suppose that the person gets one point of credit for each correct answer and that 1/3 point is deducted for each incorrect answer. What is the expected value of the person's score for the three questions?